**XBRL y AI**

Publicado por:   [Jan-Paul van der Velden](https://info.vismaconnect.nl/blog/xbrl-and-ai) el 08/04/2020

XBRL e AI: ¿un mal matrimonio o una pareja hecha en el cielo?

XBRL es ahora el estándar para informes financieros. Cuando los informes solían archivarse como documentos PDF u hojas de Excel, el estándar XBRL ha proporcionado a los reguladores datos legibles por máquina. El hecho de que sea legible por máquina ha mejorado significativamente la calidad de los datos, ya que las reglas comerciales se pueden aplicar cuando se envían los datos. Muchos reguladores quieren dar un paso más y aplicar inteligencia artificial a los datos que reciben.

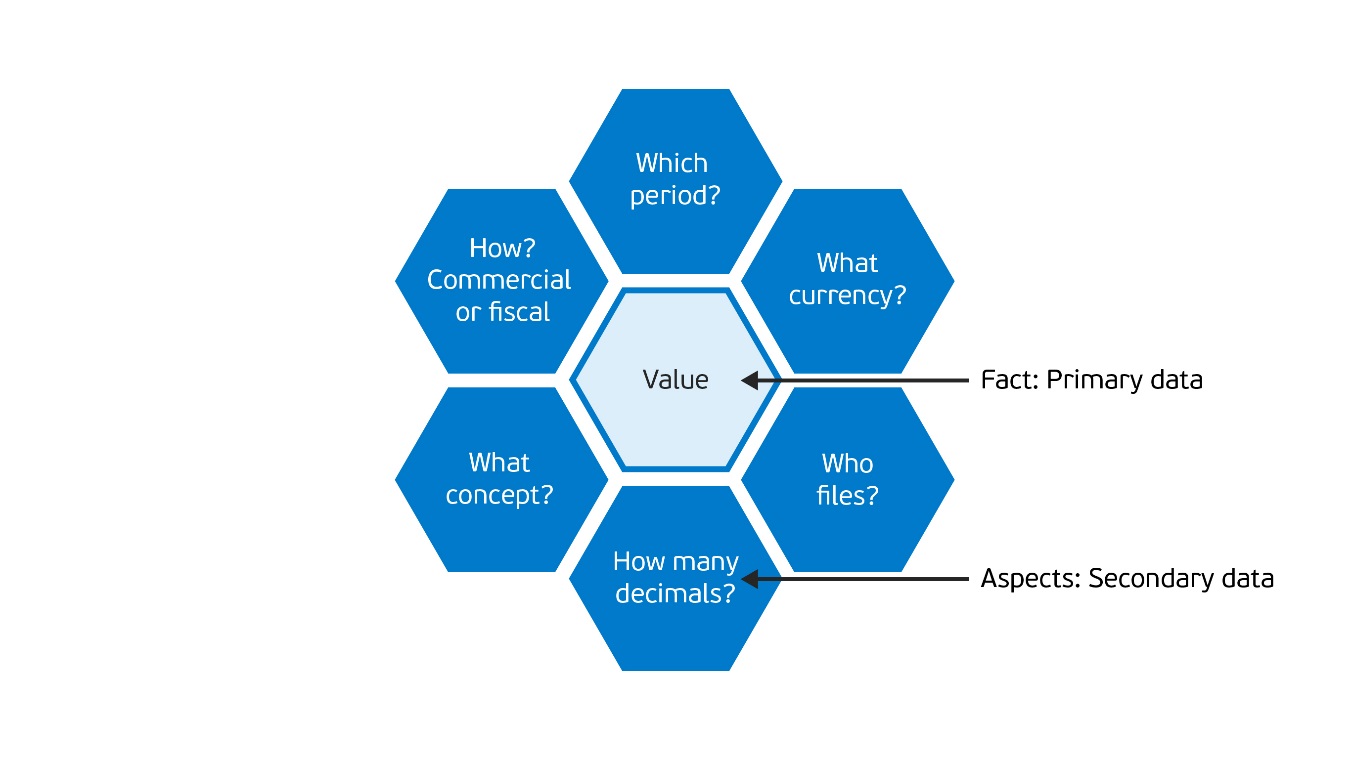
De los datos a los conocimientos

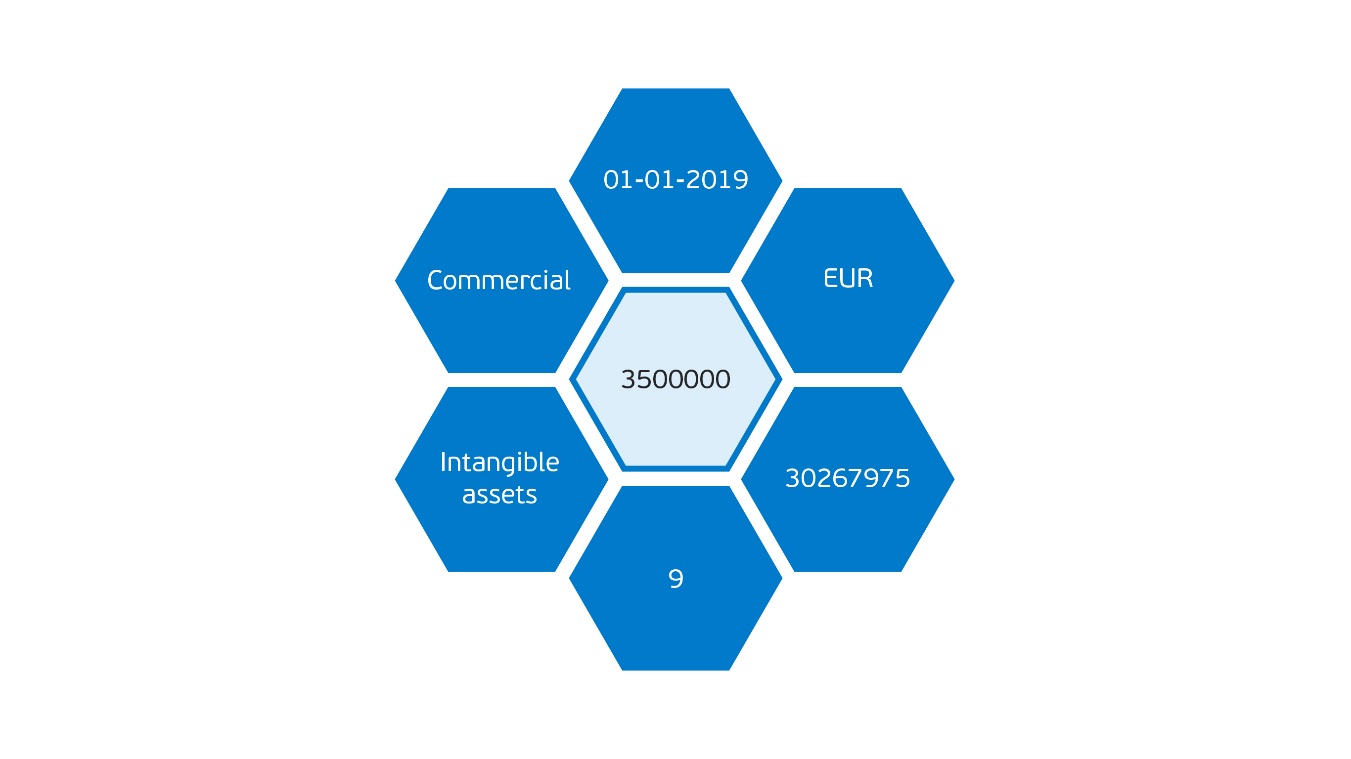
Una vez que los reguladores reciben los datos, quieren analizarlos y obtener información mediante la realización de comparaciones entre organizaciones, año tras año. Sin embargo, esta no es una tarea trivial. XBRL y AI tienen enfoques de datos muy diferentes. Para obtener una mejor comprensión, exploremos las diferencias a continuación.

* Los datos XBRL están muy estructurados. Una taxonomía XBRL describe las estructuras de datos y cómo se presentarán los datos. Contiene definiciones de reglas comerciales para salvaguardar la calidad de los datos.
* La IA se ocupa de datos que a menudo no están estructurados, con poco conocimiento de los tipos de datos y sin ningún conocimiento de las reglas comerciales.
* XBRL usa un dialecto XML, que permite estructuras anidadas.
* La IA utiliza principalmente lenguajes como Python y R, que funcionan mejor con estructuras planas (CSV y otros datos basados ​​en filas) o JSON.
* Los datos XBRL son altamente dimensionales [1](https://info.vismaconnect.nl/blog/xbrl-and-ai#Notes-1)
* En IA y aprendizaje automático, tener menos dimensiones simplifica procesos como el entrenamiento y la validación.
* XBRL tiene reglas comerciales para validar la consistencia interna de una presentación y rechaza una presentación si fallan las reglas comerciales. Esto da como resultado una mayor calidad de los datos.
* La IA no asume por adelantado una alta calidad de los datos; hay formas de manejar los datos faltantes (por ejemplo, la imputación).

Un punto de datos XBRL no es como un valor único para una columna en un archivo CSV. Un punto de datos es un valor con un grupo de aspectos que describen ese valor. Siempre que informa un valor, también informa el qué, quién y cuándo. Es decir, **qué** concepto se informa, **quién lo** informó, **cuándo** se informa.

La siguiente imagen muestra la estructura de un punto de datos XBRL [2](https://info.vismaconnect.nl/blog/xbrl-and-ai#Notes-1) :





Los datos primarios y secundarios son difíciles de manejar en técnicas de IA, como el aprendizaje automático. El aprendizaje automático funciona mejor con datos tabulares (filas y columnas).

Si queremos aplicar IA a los datos XBRL, debemos superar estos obstáculos. En Visma Connect, creemos que es muy posible aplicar IA a los datos XBRL siguiendo una serie de pasos.

Pasos de procesamiento

Para transformar los datos XBRL a un formato que sea adecuado para entornos de IA, debemos hacer lo siguiente:

Filtración

Filtre XBRL realizando una selección en ciertos aspectos como concepto, período, moneda, etc. Este paso debe realizarse utilizando un software que tenga conocimiento de XBRL. El software debe saber cómo manejar las estructuras de datos que usa XBRL.

Ejemplo:

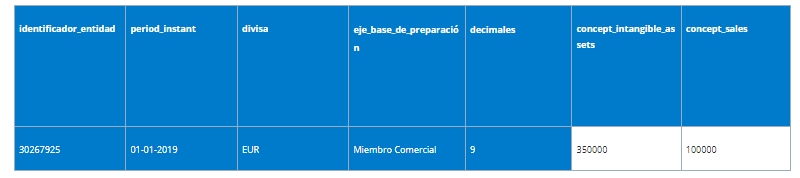
* Filtrar por identificación de entidad 30267925.
* Filtrar solo por moneda Euro.
* Etc.

Aplastamiento

Una vez que hayamos filtrado los datos y sepamos qué datos tenemos en nuestras manos, podemos aplanar los datos. En lugar de informar la moneda con cada punto de datos, podemos extraer el aspecto de la moneda y convertirlo en un punto de datos separado. Lo mismo ocurre con otros aspectos como el período.

El siguiente paso es combinar puntos de datos en una fila, lo que significa que todos los puntos de datos que informan de manera similar se colocan en una fila. Estos datos tabulares son más fáciles de administrar en lenguajes como R y Python.

El punto de datos XBRL en la imagen de muestra de arriba se verá así después de aplicar estos pasos de procesamiento:



Tenga en cuenta que el concepto Ventas también se informa en la misma fila, que tiene los mismos datos secundarios que el concepto de Activos intangibles.

Análisis de los datos

Una vez que hemos filtrado y compactado los datos, conservamos la alta calidad de los datos y la estructura descrita en la taxonomía.

A medida que analizamos los datos, tenemos que eliminar algunos campos que no juegan un papel significativo en el entrenamiento del modelo. Esto dará como resultado un modelo de mejor rendimiento. Este proceso se llama reducción de dimensión.

Muchos proyectos de IA fallan debido a la mala calidad de los datos o los resultados son deficientes porque la estructura de los datos no está bien descrita. Los ingenieros de datos normalmente tienen dificultades para procesar datos de tan mala calidad. En los proyectos de IA, la mayor parte del tiempo se dedica al procesamiento de datos, no al análisis de datos en sí. Aquí, el uso de datos XBRL estructurados y de alta calidad resulta ser una ventaja.

¿Quieres comenzar tu viaje de IA?

En Visma Connect, tenemos el conocimiento y las herramientas para trabajar con XBRL. Creamos taxonomías para muchos de nuestros clientes en el ecosistema SBR. Los sistemas de Visma Connect validan los mensajes procesados ​​por la plataforma y entregan los mensajes a varias organizaciones.

Vemos que procesar datos XBRL puede resultar difícil para las organizaciones que los reciben. Visma Connect implementó la especificación XBRL OIM para transformar CSV y JSON en XBRL y viceversa. Esto abre una forma de aplicar IA en esos datos.

Los casos de uso relacionados con la IA son numerosos. La detección de fraudes es, en esencia, un problema de clasificación. AI puede separar transacciones auténticas y fraudulentas utilizando los puntos de datos proporcionados. La oficina de impuestos podría hacer esto utilizando datos históricos. Si eso no está disponible, el aprendizaje no supervisado podría aplicarse para agrupar transacciones similares, formar grupos distintos e identificar claramente valores atípicos. NLP (procesamiento del lenguaje natural) se puede aplicar a los textos de divulgación.

Nuestro equipo de inteligencia artificial tiene experiencia con todo tipo de datos, incluidos los datos XBRL. Podemos trabajar junto con nuestros clientes para materializar la promesa de la IA.

Contacto



Contáctenos para conocer más sobre las posibilidades de XBRL y AI en su organización.

Notas

1: La definición del término dimensión en XBRL es diferente a la de AI. En aras de la simplicidad, usamos una y la misma palabra aquí.

2: En aras de la simplicidad, se omitieron algunos detalles de las imágenes.